Enhanced Weed Detection for Cotton Crops using YOLO11 and Soft-NMS

Anvith RH[1], Raghavarshini K[2], Kadappa S[3], Daneshwari M[4], Shashank Hegade[5], and Sneha Varur[6]

School of Computer Science and Engineering
KLE Technological University, Hubballi, Karnataka, India
01fe22bcs192@kletech.ac.in, 01fe22bcs101@kletech.ac.in, 01fe22bcs181@kletech.ac.in,
01fe22bcs253@kletech.ac.in, shashank.hegade@kletech.ac.in, sneha.varur@kletech.ac.in

Abstract. Weed management is a critical challenge in agriculture, as uncontrolled weed growth can significantly reduce crop yields and increase production costs. This study demonstrates the application of YOLO (You Only Look Once), a highly efficient and accurate real-time object detection framework. The proposed method uses the CottonWeedDet3 dataset, comprising 848 images of three weed classes—carpetweed, morningglory, and palmer amaranth—for benchmarking. In our weed detection pipeline, we observed that YOLO11 may suppress overlapping bounding boxes with low confidence scores, potentially leading to false negatives. To address this, we integrated a post-processing layer known as Soft-NMS (Soft Non-Maximum Suppression) into the pipeline, applied after the YOLO11 model's training phase. Unlike traditional NMS, Soft-NMS reduces the confidence scores of overlapping boxes instead of discarding them outright. The YOLO11 model achieved an overall mean Average Precision (mAP) of 0.969, with class-wise mAP scores of 0.949 for carpetweed, 0.970 for morningglory, and 0.989 for palmer amaranth. In comparison, YOLOv10 achieved an mAP of 0.916 on the same dataset, highlighting the performance improvements of the proposed method. Future work will focus on replacing bottleneck modules with Dilation-wise Residual Modules (DWR) and adding Multi-Scale Modules (MSBlock) to improve multi-scale feature capture and contextual understanding.

Keywords: Weed Detection, YOLO11(You Only Look Once), Soft Non-Maximum Suppression (Soft NMS),mAP(mean average precision).

1 Introduction

Cotton is a crop often termed as "White Gold" because of its significant role in global agriculture. Among several challenges in its cultivation, weed competition is one of the most significant. Weeds impose biotic stress on the cotton crop by competing for essential resources such as water, nutrients, and sunlight. This competition results in a drastic reduction in crop yields, with weed stress contributing to an estimated 43% of global crop yield loss [16, 10]. Furthermore, weeds act as vectors for pests and diseases, exacerbating the damage. To mitigate these challenges, there is a pressing need to develop precise and efficient weed detection technologies to ensure the safety and productivity of cotton farming.

Conventional weed control approaches, including mechanical methods like plowing and mowing and chemical methods using herbicides, have been widely employed. However, these methods are fraught with limitations, such as resistant weed populations, pesticide residues, and loss of biodiversity, emphasizing the need for non-chemical alternatives [18]. Recent advancements in computer vision and machine learning, particularly object detection models like YOLO, offer promising potential for automating weed detection and management in real-time[4], enabling targeted interventions that reduce herbicide use, lower costs, and enhance crop yield and quality [3].

The optimization of YOLO for cotton weed detection focuses on enhancing precision and effectiveness in identifying weeds in cotton farming areas. It leverages deep learning and computer vision to provide scalable solutions for weed management, aligning with the principles of precision agriculture. With the evolution of precision agriculture, the integration of advanced technologies like YOLO represents a critical step toward more sustainable, cost-effective, and environmentally friendly farming practices [19], [3].

The latest development in the YOLO series, YOLO11, unveiled at the YOLO Vision 2024 (YV24) conference, marks a significant leap forward in real-time object detection technology[13]. Building upon the

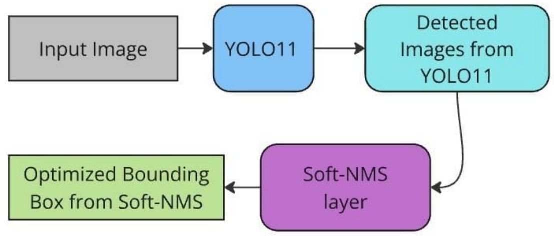

Fig. 1. Model Architecture Overview showing YOLO11 Object Detection Followed by Soft-NMS for Bounding Box Optimization

foundation established by YOLOv1[9], YOLO11 introduces substantial enhancements in both architecture and training methodologies, pushing the boundaries of accuracy, speed, and efficiency[9].

In reference to Fig. 1 our proposed architecture, a Soft-NMS layer has been introduced as a post-processing step to optimize overlapping bounding boxes. Input images are first processed by the YOLOv11 model for weed detection, generating bounding boxes around the detected weeds. These outputs are then passed to the Soft-NMS layer, which aims to reduce the overlap between bounding boxes while preserving the most relevant detections. The key feature of the Soft-NMS layer is its ability to assign lower scores to overlapping boxes, ensuring that the most accurate bounding boxes are retained. Finally, the output consists of optimized images with refined bounding boxes for improved weed localization.

The experimental results demonstrate the effectiveness of the proposed approach, showing a significant improvement in mean Average Precision (mAP), increasing from 0.911 with YOLOv10 to 0.969 with YOLO11. The integration of the Soft Non-Maximum Suppression (Soft NMS) layer was key to improving detection reliability, particularly in scenarios with overlapping bounding boxes. Soft NMS helped to preserve true positives while reducing false positives, ensuring more accurate weed localization. This enhancement is especially critical in dense fields where weeds and crops may appear similar.

The paper is organized as follows: Section 2 reviews related literature on weed detection, focusing on YOLO-based methods. Section 3 explains the proposed methodology, including YOLO11 enhancements and Soft NMS integration. Section 4 presents the results, concluding with findings and future work suggestions.

2 Related Work

Advancements in weed detection algorithms have significantly transformed precision agriculture, enabling efficient and automated solutions for crop management. Conventional weed control approaches were mechanized and chemical in nature. The most common approach to weed control is mechanical, such as plowing and mowing, while chemical control relies on herbicides. However, these methods have numerous shortcomings. Overreliance on herbicides leads to resistant weed populations and risks such as pesticide residues and loss of biodiversity [18]. Resistance has become a significant challenge with the rapid development of herbicide-resistant weeds, emphasizing the need for alternative non-chemical weed control strategies [18].

Object detection frameworks are generally classified into two categories: two-stage and one-stage models. Two-stage models like Faster R-CNN excel in precise localization, extracting features before classifying objects, whereas one-stage models like YOLO directly predict object classes and locations, making them suitable for real-time applications due to their faster inference times [16][13].

Numerous studies have explored the applicability of YOLO in agricultural settings. For instance, Dang et al. proposed the YOLO Weeds benchmark for detecting multi-class weeds in cotton fields, achieving significant advancements in detection precision [5]. Similarly, Pérez-Porras et al. demonstrated the efficiency of YOLOv5s for detecting Papaver rhoeas in winter wheat fields[6], attaining an F1-score of 75.3% and mAP of 76.2%, highlighting the importance of site-specific weed management strategies [14].

Innovations in YOLO architectures continue to enhance their versatility[17]. Alif et al. reviewed YOLO variants from YOLOv1 to YOLOv10, detailing their evolution and applications in agricultural domains [2]. Zhang et al. proposed a lightweight weed detection model incorporating contextual information fusion, which improved the detection of small and overlapping targets [18]. Shao et al. introduced GTCBS-YOLOv5s, capable of identifying six weed types in rice paddies under varying lighting conditions[11].

In cotton crops, Zheng et al. improved YOLOv8 for weed recognition by optimizing its detection algorithm, achieving superior results in precision and robustness [20]. Sharma et al. compared the performance of YOLOv8 through YOLO11 and Faster R-CNN, demonstrating that YOLO11 achieved the highest mAP of 0.969 with enhanced speed and accuracy [16]. Zhou et al. applied YOLO11 for intelligent object segmentation at construction sites, showcasing its adaptability and precision across diverse domains, further emphasizing its potential in agricultural applications [8].

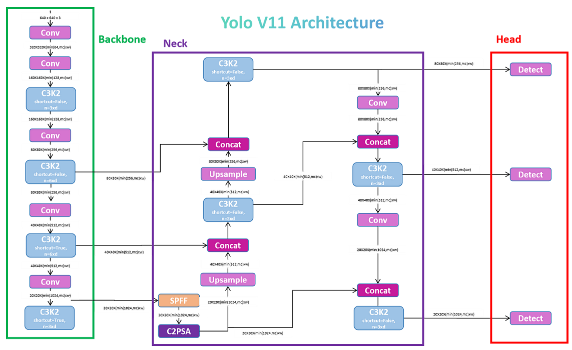

Fig. 2. Architecture of the YOLO11 model, showing the backbone, neck, and head components.

Future research focuses on further optimizing YOLO models by integrating hybrid frameworks, leveraging transfer learning, and exploring edge-device deployment with multispectral data [1]. Such advancements are poised to revolutionize precision agriculture, enhancing weed detection and crop monitoring [3].

Figure 2 illustrates the architecture of the YOLO11 object detection model. It is composed of three main sections: a backbone network for feature extraction, a neck network for feature aggregation, and a head network for object detection. Key components and their parameters are indicated.

In this paper, we explore the application of the YOLO11 model for developing a weed detection system specifically for cotton crop weeds, with the model trained to identify three different weed classes—carpetweed, morningglory, and palmer amaranth—using the CottonWeedDet3 dataset, which consists of 848 images. To improve the accuracy of weed detection, especially in challenging scenarios where weeds are densely packed, we enhance the original YOLO11 architecture by replacing the standard Non-Maximum Suppression (NMS) layer in the backbone with Soft-NMS. This modification allows the model to handle overlapping bounding boxes more effectively, reducing the suppression of boxes with high overlap but lower scores, ultimately leading to more precise detection in complex field conditions. The integration of Soft-NMS addresses key limitations of standalone YOLO models, which often discard overlapping boxes outright, resulting in false negatives and lower reliability in dense environments. By assigning lower confidence scores to overlapping boxes instead of discarding them, Soft-NMS preserves true positives, ensuring more accurate localization.

3 Methodology

In this paper, the proposed Weed Detection model provides an edge over strictly YOLO-dependent models by integrating with Soft Non-Maximum Suppression (Soft-NMS). It addresses the challenges of detecting overlapping weeds in agricultural fields. The following methodology includes model selection, dataset preparation, training, post-processing, and validation, each tailored to optimize the detection of weeds in diverse and challenging conditions.

3.1 Model Selection and Motivation

The model best fit for the given weed detection task was YOLO11 due to its enhanced speed and accuracy for object detection tasks[17]. Its ability to efficiently handle dense objects of varying sizes and shapes makes it ideal for agricultural applications [15]. Additionally, we replaced the default NMS block in the YOLO11 architecture with a Soft-NMS block, which improves detection in dense weed clusters by avoiding overlapping boxes.

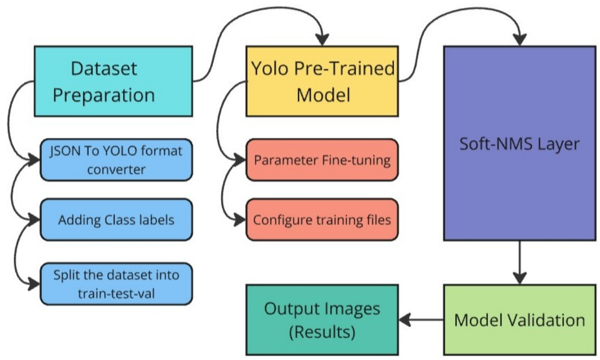

Fig. 3. Workflow Diagram Illustrating the Model Training Process, from Dataset Preparation and Pre-trained YOLO Model Fine-tuning to Soft-NMS Application and Final Model Validation.

The workflow (Fig. 3) shows dataset preparation, YOLO model fine-tuning, and integration of a Soft-NMS layer. It improves detection reliability and outputs validated results, the model consists of three primary phases: dataset preparation, YOLO11 fine-tuning, and integration of the Soft-NMS layer (Fig. 3).

3.2 Implementation

The integrated YOLO11 model, combined with a Soft Non-Maximum Suppression (SoftNMS) layer, is used to improve weed detection accuracy by addressing overlapping bounding boxes and minimizing false positives. This integrated approach is employed to effectively detect the three weed species (carpetweed, morning-glory, and palmer amaranth) in dense agricultural environments, enhancing the model's generalization and performance on the cotton weed dataset.

Dataset Preparation Dataset preparation involved sourcing a diverse dataset with multiple classes of weeds. Each image was labeled using the YOLO format refer equation 1, which specifies bounding boxes as normalized values relative to image dimensions.

$$x_{\text{center}} = \frac{x_1 + x_2}{2W}, \quad y_{\text{center}} = \frac{y_1 + y_2}{2H}[7] \tag{1}$$

Here, x_1, x_2, y_1, y_2 are the bounding box coordinates, and W, H are the image width and height. The dataset was divided into training (80%), validation (10%), and testing (10%) subsets. To improve generalization, various augmentations such as flipping, rotation, scaling, and contrast adjustments were applied.

Model Training The YOLO11 model training began with pre-trained weights, reducing training time and improving initial accuracy. The training process optimized three key components: bounding box loss, confidence loss, and class loss[12]. The total loss is given in equation 2

$$L = L_{\text{box}} + \lambda_{\text{conf}} L_{\text{conf}} + \lambda_{\text{cls}} L_{\text{cls}}[7] \tag{2}$$

where λ_{conf} and λ_{cls} are weighting factors. Bounding Box Loss, ensures accurate localization of weeds. Confidence Loss, differentiates between object and background regions. Class Loss, ensures accurate classification of weed types.

The YOLO11 model works by dividing the image into patches, with each patch predicting bounding boxes and associated confidence scores. The model predicts a set of class probabilities and a confidence score for each box, reflecting the probability of containing an object. The objective is to minimize the following loss function given in equation 3 [7]

$$
\begin{aligned}
L_{\text{total}} = \lambda_{\text{coord}} \sum_i & \left[(x_i - \hat{x_i})^2 + (y_i - \hat{y_i})^2 + (w_i - \hat{w_i})^2 \right. \\
& \left. +(h_i - \hat{h_i})^2 \right] + \lambda_{\text{obj}} \sum_i \left[(C_i - \hat{C_i})^2 \right] \\
& + \lambda_{\text{noobj}} \sum_i \left[(1 - \hat{C_i})^2 \right]
\end{aligned} \tag{3}
$$

In this equation: x_i, y_i, w_i, h_i are the predicted bounding box parameters (center coordinates, width, height), $\hat{x_i}, \hat{y_i}, \hat{w_i}, \hat{h_i}$ are the true bounding box parameters, C_i is the predicted object confidence, $\hat{C_i}$ is the true object confidence.

Training was carried out locally on a system equipped with an NVIDIA RTX 3050 GPU (6 GB VRAM) and 16 GB RAM. This configuration provided sufficient computational resources for training the model and conducting test runs on unseen data. Hyperparameter tuning was a critical step, involving experimental runs with varying epochs (50, 75, 100), batch sizes (16, 32), and learning rates (0.001, 0.0005). This tuning process identified the optimal configuration that balanced the model's accuracy and efficiency.

Soft-NMS Integration Soft-NMS is a part of the post-processing layer, which primarily comes into play during inference. It refines bounding box outputs. Unlike traditional NMS, which suppresses overlapping bounding boxes, Soft-NMS reduces the confidence scores of overlapping bounding boxes using a decaying function based on the IoU,detailed description being shown in algorithm 1 [5] and algorithm 2 [13] and the confidence scores can be calculated as per the equation 4

$$
s_i = s_i \times (1 - \text{IoU})^{\alpha} [20] \tag{4}
$$

Here, α is a decay parameter. This adjustment allows the model to retain high-confidence predictions even in cases of significant overlap, which is particularly common in dense weed clusters.

height.8pt depth0pt
Algorithm 1 ComputeIoU Algorithm
pt

procedure ComputeIoU(Box1, Box2)
    Calculate the intersection area of `Box1` and `Box2`
    Calculate the union area of `Box1` and `Box2`
    return IoU = $\frac{\text{intersection}}{\text{union}}$ $\frac{\text{3em}}{\text{3em}}$
end procedure

height.8pt depth0pt
Algorithm 2 Soft-NMS Algorithm
pt

procedure Soft-NMS(Boxes, Scores)
  Find the box $b_{\max}$ with the highest score
  for each bounding box b_i **do**
    for each remaining box b_j **do**
      Compute IoU$(b_{\max}, b_j)$
      if IoU > threshold **then**
        Update the score of b_j using:

new 3em score$_j$ = c

      end if
    end for
    Keep boxes with updated scores above a predefined threshold
  end for
  return Filtered bounding boxes and their scores
end procedure

[20]

Soft-NMS aids in retaining bounding boxes in dense environments and reduces false negatives. However, the improved accuracy comes at the cost of computational speed when working with large datasets, as it introduces additional calculations for redefining the bounding boxes[20].

3.3 Validation and Testing

Validation of the model used metrics such as mean Average Precision (mAP). mAP@0.5 evaluates precision and recall at an IoU threshold of 0.5, while mAP@0.5:0.95 averages these metrics across IoU thresholds ranging from 0.5 to 0.95. The model's performance was further tested on an independent dataset to ensure its generalizability and robustness. Qualitative analysis included visualizations of detected bounding boxes, which confirmed the model's ability to accurately detect overlapping and densely packed weeds in complex field scenarios.

Additionally, the recall curve for various weed classes, as shown in Figure 2, highlights the model's ability to detect different weed species effectively. The recall values for specific classes include carpetweed (0.949), morning glory (0.97), and Palmer amaranth (0.989), with an overall recall of 0.969 across all classes.

4 Results and Discussion

1) Dataset Description: The CottonWeedDet3 dataset contains 848 RGB images in JPEG (`.jpg`) format. These images are captured with a resolution suitable for weed detection tasks, typically in the range of moderate resolution (e.g., 640x480 or higher). The dataset includes images of three weed species—carpetweed, morningglory, and palmer amaranth—commonly found in cotton fields under natural field light conditions. Additionally, each image is annotated with bounding boxes to highlight the presence of these weeds. The bounding box annotations are stored in JSON format.

2) Results: The class-specific performance metrics of YOLO11 are presented in Table 2. These metrics show that the model achieved exceptionally high precision and mAP across diverse weed classes.

The model's ability to generalize well was further validated with visual representations of Precision-Recall (PR) curves and F1 curves during training. The Precision-Recall curve, as shown in Fig. 4, illustrates the recall trade-off for different precision levels. It demonstrates the outstanding performance of YOLO11 throughout training. The model exhibited exceptional accuracy in detecting weeds, particularly in distinguishing different weed classes, as shown in Fig. 5. This includes dense clusters of overlapping weeds where accurate detection is most challenging.

Table 1. Overall Performance Metrics of YOLO11

Metric	YOLO11
Precision (P)	0.954
Recall (R)	0.886
mAP@0.5	0.969
mAP@0.5:0.95	0.938

Table 1 highlights the overall performance metrics of YOLO11, demonstrating its precision and robustness in weed detection tasks.

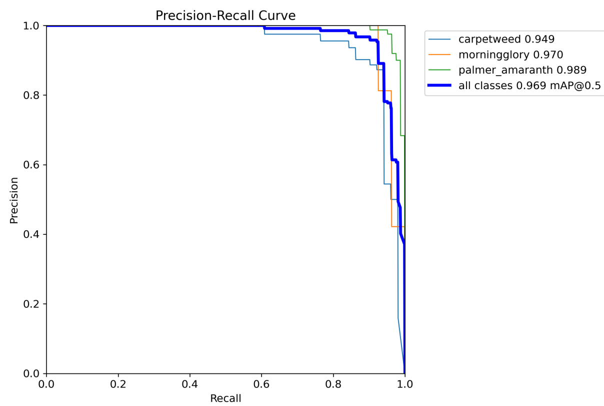

Fig. 4. Precision-recall Curve of YOLO11 During Training, Demonstrating High Precision and Recall Across Diverse Weed Classes.

Table 2. Performance Metrics for Different Weed Classes

Class	Precision	Recall	mAP@0.5	mAP@0.5:0.95
Carpetweed	0.926	0.863	0.949	0.911
Morningglory	0.936	0.926	0.970	0.952
Palmer Amaranth	1.000	0.870	0.989	0.950

4.1 Soft-NMS Integration and Bounding Box Adjustments

To further enhance YOLO11's performance in detecting dense and overlapping weed clusters, Soft-NMS (Non-Maximum Suppression) was implemented during post-processing. The advantages of Soft-NMS over traditional NMS include retaining overlapping boxes with reduced confidence scores while reducing false negatives.

Figures 6 and 7 show the qualitative improvement offered by Soft-NMS. Specifically, detection types before Soft-NMS implementation are displayed in Fig. 6, while the refined results after its integration are demonstrated in Fig. 7.

Key improvements achieved through Soft-NMS include precise localization of dense weeds, improved detection of partially occluded and small weeds, and higher overall confidence in overlapping scenarios, comparative test was conducted using YOLOv10 to analyse the improvement of YOLO11 model with Soft-NMS , the results showed a clear improvement in our models accuracy over YOLOv10.

Fig. 5. Examples of Weed Detection By the YOLO11 Model.

Fig. 6. Output of Weed Detection From YOLO11 Model.

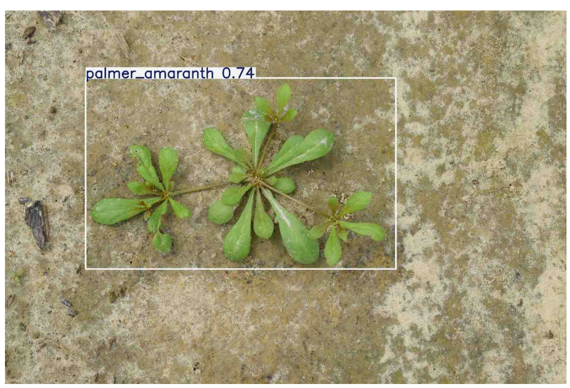

Fig. 7. Impact of Soft Non-Maximum Suppression (Soft-NMS) on YOLO11 Detection Results.

The comparative performance of YOLOv10 and YOLO11 is presented in Table 3, underscoring the significant gains achieved in precision, recall, and mAP due to Soft-NMS and architectural advancements.

The YOLO11 model demonstrates robust generalizability and adaptability across diverse contexts, reinforcing its potential for precision agriculture applications. For example, its ability to address repeated detections in dense clusters (e.g., Carpetweed), missed small targets (e.g., Morningglory), and overlapping weeds (e.g., Palmer Amaranth) makes it a highly effective tool for weed detection in real-world farming scenarios.

Table 3. Comparison of YOLO11 and YOLOv10 performance metrics.

Metric	YOLO11	YOLOv10
Precision (P)	0.954	0.900
Recall (R)	0.886	0.851
mAP@0.5	0.969	0.916
mAP@0.5:0.95	0.938	0.878

5 Conclusion

This study demonstrates the effectiveness of YOLO11 with Soft Non-Maximum Suppression (Soft NMS) for weed detection in cotton crops, achieving an overall mAP of 0.969 compared to 0.911 in YOLO v10. The introduction of the Soft NMS layer proved instrumental in preserving bounding boxes, leading to better interpretation and localization of detected weeds. These results highlight the robustness of our approach and its potential for precision agriculture. The Soft-NMS layer over YOLO11 improves detection accuracy but adds computational overhead, slowing inference times. This extra calculation for modifying confidence scores increases processing time, especially with large datasets. Additionally, it raises memory usage, impacting performance on resource-limited devices. In real-time applications, this tradeoff between accuracy and speed may limit its practicality.

However, there is scope for further improvement. Replacing the BottleNeck modules in the C2f network with Dilation-wise Residual Modules (DWR) could enhance the model's ability to capture multi-scale features. Adding a Multi-Scale Module (MSBlock) in the final backbone layer may provide better contextual understanding by leveraging features at various resolutions. Additionally, integrating the Adaptively Spatial Feature Fusion (ASFF) mechanism could resolve spatial inconsistencies during feature fusion, enhancing overall detection accuracy. These advancements would make the model even more robust and effective, paving the way for improved weed management practices in diverse farming conditions.

References

[1] O.G. Ajayi, J. Ashi, and B. Guda. "Performance evaluation of YOLO v5 model for automatic crop and weed classification on UAV images". In: *Smart Agricultural Technology* 5 (2023).

[2] M.A.R. Alif et al. "YOLOv1 to YOLOv10: A Comprehensive Review of YOLO Variants and Their Application in the Agricultural Domain". In: (2024).

[3] Various Authors. "Intelligent Weeding Solutions for Cotton". In: *Plants* 13 (2024).

[4] B. Cui et al. "Exploring the YOLO-FT Deep Learning Algorithm for UAV-Based Smart Agriculture Detection in Communication Networks". In: *IEEE Transactions on Network and Service Management* (2024).

[5] F. Dang et al. "YOLOWeeds: a novel benchmark of YOLO object detectors for multi-class weed detection in cotton production systems". In: *Computers and Electronics in Agriculture* 205 (2023).

[6] T. Delleji et al. "An upgraded-yolo with object augmentation: Mini-uav detection under low-visibility conditions by improving deep neural networks". In: *Operations Research Forum* 3.4 (2022), p. 60.

[7] W. Dong. "Large-scale Remote Sensing Image Target Recognition and Automatic Annotation". In: *arXiv preprint arXiv:2411.07802* (2024).

[8] L. He et al. "Research and Application of YOLO11-Based Object Segmentation in Intelligent Recognition at Construction Sites". In: *Buildings* 14.12 (2024), p. 3777.

[9] R. Khanam and M. Hussain. "YOLO11: An Overview of the Key Architectural Enhancements". In: *arXiv preprint arXiv:2410.17725* (2024). DOI: 10.48550/arXiv.2410.17725.

[10] D. Li et al. "YOLO-JD: A Deep Learning Network for Jute Diseases and Pests Detection from Images". In: *Plants* 11.7 (2022), Article 937. DOI: 10.3390/plants11070937.

[11] Z.E. Maretanio et al. "Estimation of Rice Field Area Using YOLO Method to Support Smart Agriculture System". In: *2024 International Electronics Symposium (IES)*. IEEE, 2024, pp. 562–568.

[12] Y. Natij et al. "Evaluating the Performance of YOLO Object Detectors for Plant Disease Detection". In: *2024 11th International Conference on Wireless Networks and Mobile Communications (WINCOM)*. IEEE, 2024, pp. 1–6.

[13] P.K. Pativada. "Real-time detection and classification of plant seeds using YOLOv8 object detection model". In: (2024).

[14] F.J. Pérez-Porras et al. "Early and on-ground image-based detection of poppy (Papaver rhoeas) in wheat using YOLO architectures". In: *Weed Science* 71 (2023), pp. 50–58.

[15] D.C. Rodríguez-Lira et al. "Comparative Analysis of YOLO Models for Bean Leaf Disease Detection in Natural Environments". In: *AgriEngineering* 6.4 (2024), pp. 4585–4603.

[16] A. Sharma, V. Kumar, and L. Longchamps. "Comparative performance of YOLOv8, YOLOv9, YOLOv10, YOLOv11 and Faster R-CNN models for detection of multiple weed species". In: *Smart Agricultural Technology* 9 (2024).

[17] J. Wei et al. "GFS-YOLO11: A Maturity Detection Model for Multi-Variety Tomato". In: *Agronomy* 14.11 (2024), p. 2644.

[18] C. Zhang, J. Liu, H. Li, et al. "Weed Detection Method Based on Lightweight and Contextual Information Fusion". In: *Applied Sciences* 13.24 (2023).

[19] H. Zhang, Z. Wang, Y. Guo, et al. "Weed Detection in Peanut Fields Based on Machine Vision". In: *Agriculture* 12.1541 (2022).

[20] L. Zheng et al. "Improvement of the YOLOv8 Model in the Optimization of the Weed Recognition Algorithm in Cotton Field". In: *Plants* 13.13 (2024), p. 1843.